

The Lean Thinking Machine: The Romanian Context

Author: Nicolaie Constantinescu, Information Architect, www.kosson.ro

kosson@gmail.com

Abstract

When a new idea arrives in a manifested form it is encompassing the whole set of information environment parameters that triggered that very moment. It is because you are able to discover, order, manipulate and refactor the raw data sets in your own knowledge set. A quite unique set. Many components of the set are manifested augmentations of one unique set of research habits. What happens when in the process of shaping the original idea actor like LAM specialists, software and hardware tools, legal framework and social graph are involved proposing different and more efficient approaches? Some answers arrive when you deconstruct the idea's background, as personal knowledge background, libraries, cultural frameworks and technological means are seen as real fairy godmothers of the future discoveries. But most of the time it is all about hacking your own way into the field being into a continuous competition with time and different limitations. The study will offer some paths using as many augmentation skills and tools as the aforementioned actors may offer in order to become an elegant lean thinking machine.

Keywords: data, lams, Romanian electronic resources, meshups, data meshups

Introduction

It is told that we have stepped into a new era: Anthropocene¹ – a period in which man has the capacity to influence many aspects governing his existence mastering his natural environment. Many are attracted by the “knowledge society” buzzword looking for signs of improvement for the entire information ecosystem in which they perform. It is easy to observe that such an enabling informational for reaching knowledge it is a much-desired goal but is an elusive target concerning the understanding and supporting of the infrastructures, let that be human and technical. There is also a nascent gap between those who understand and develop technologies and those who are seen as mere consumers. A solution to this trend is to have a middle layer of specialists to act as enablers to the continuous enriching space of information. These actors are needed to give a chance to the less technology savvies to gain knowledge benefits from scores of resources growing out there.

A 10.000 feet perspective on the resources

The libraries

On Romanian information brokering scene there is still a valuable good traded, and at times forgotten for its use and purposes being treasured for what it stands: the book. Libraries are concentrating their

¹ <http://en.wikipedia.org/wiki/Anthropocene>

efforts on the services dedicated to the traditional support for the information. Some of the public libraries have managed to pull their resources together in terms of federated search: <http://catalog.biblio.ro/opac/advancedsearch> (TinREAD) and the Romanian Library Network Science & Technology² (METALIB) catalogue: <http://aleph.edu.ro/>.

There is another function growing in importance lately at least on public library level: cultural animation. This trend is a direct competitor speaking in terms of time resources allocation with the primary mission of the memory institutions: delivering quality information and access to fine quality electronic resources. This is not to be taken as a negative influence, but as a component that needs further fine-tuning not to loose the prime mission, but to realize a harmoniously articulated structure.

The latest comer is digitization³. The efforts of bringing Romanian heritage online⁴ are on the same level with those of ensuring a large access to the most visible editor's databases out there⁵: continuous development based on networking capacities of some institutions on the background of governmental funding stagnation for a national framework.

Unfortunately Romania lags behind the other European states according to the eGovernment Benchmark Survey 2010⁶ (published on February 2011).

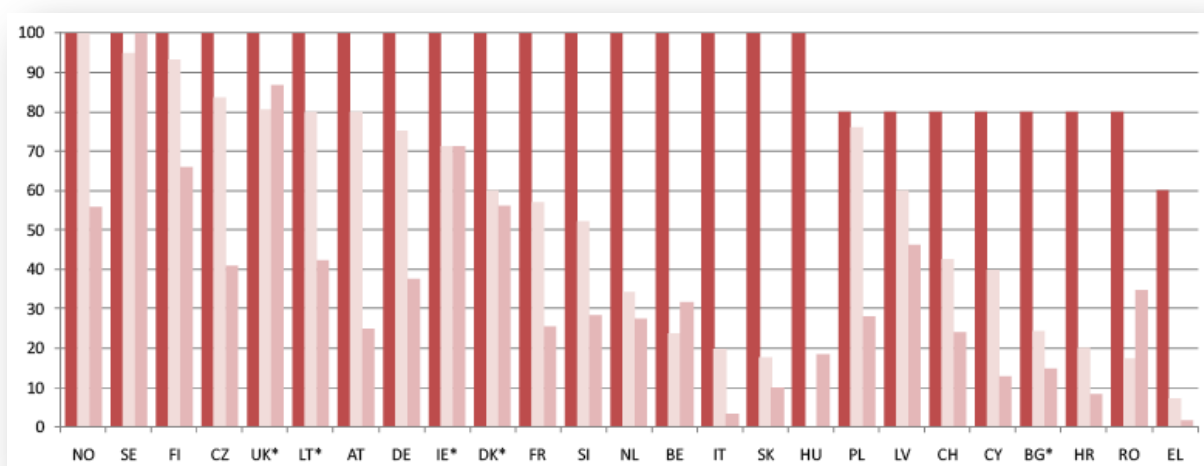


Figure 1: Public libraries – sophistication at levels: country, NUTS 5a and 4 or 5b (in %)

It is not surprising to find on figures that the main cities (NUTS5a) are lagging behind (17) the rural areas (NUTS 4) marking a higher score (35).

² http://nusidoc.upb.ro/activitati_e.htm

³ http://www.guv.ro/resursele-culturale-nationale-stocate-si-digitizate_11a103698.html

⁴ <http://www.bibnat.ro/Biblioteca-Digitala-Nationala-s135-ro.htm>

⁵ <http://www.anelis.ro/>

⁶ <http://www.epractice.eu/en/library/5283331>

Services reported at NUTS Levels 4 or 5	Proportion of countries in which available (%)																
		HU	NL	AT	PL	PT	RO	SI	SK	FI	SE	UK	HR	IS	TR	NO	CH
Public libraries	84%	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓

Figure 2: Services reported at NUTS levels 4 or 5, by type of service and by country

Overall figures say that Romania is qualifying to the Member States that score good figures, but taken in the services provided it is closing the ranks as shown in Figure 1.

Much of the efforts aimed to elevate the national library system into a coherent electronic collaboration network were concentrated around the year 2008. The concrete actions were grinded to a halt due to the worsening global economical environment and a slow reaction of all the actors in this circumstance. Looking at the big picture it is easy to see that the library system didn't gained a level of coherence in it's actions and many of the successes registered are due to local efforts of the individual institution and their mobilisation capacity in establishing projects of local interest. In spite of the budget cuts, incapacity to attract fresh personal and the continuous pressure to perform as users expected, there are islands of cooperation and real networking happening backed by external NGO funding and in some cases through individual funding gained from research funding.

The library system is in a continuous rewiring due to the fact that some of the most active stakeholders managed to see their institutions in big European projects like Europeana, Manuscriptorium, The European Library and Rediscover.

There are also good and consistent electronic collections that worth mentioning, but we will deal this aspect further down.

The archives

Most of the online easy findable output from the Romanian archives is comprised of online copyrighted inventories⁷ published as pdfs. They contain valuable data for those who are interest to localise a particular batch of records. There is another effort of bringing some of the archives treasures through the "Online Photo Collection"⁸ - a coordinated effort by the Romanian National Archives and The Institute for the investigation of Communist Crimes in Romania⁹. The latter also has digitised an important piece of Romanian communist past establishing a collection of digitised private collections encompassing Romanian exile memory.

Museums and other memory institutions

On this particular level most of the digital resources are pulled from digitisation collaborative actions between libraries and museums on local or regional level^{10 11 12}. Most of the visible efforts are

⁷ <http://www.arhivelenationale.ro/index.php?lan=0&page=313>

⁸ <http://www.arhivelenationale.ro/fototeca/?lang=en>

⁹ http://www.crimelecomunismului.ro/en/about_iiccr/institute/

¹⁰ County Library Kájoni János, Miercurea-Ciuc in cooperation with the Haáz Rezső Museum Library, Odorheiu-Secuiesc (external funding)

¹¹ Digitisation of the ecclesiastical heritage of the Ortodox Parish of Sfântu Gheorghe (government funding) <http://www.mncr.ro/evenimente%20culturale+programe%20proiecte%20si%20alte%20activitati+2011+digitizarea%20patrimoniului%20ecleziastic%20din%20cadrul%20colectiei%20parohiei%20ortodoxe%20i%20sfantu%20gheorghe.html>

aggregated by CIMEC - Institute for Cultural Memory leading the efforts on this level¹³. Being the national aggregator for Europeana project, CIMEC put together one of the most useful digital resources collections: Mobile Cultural Objects Classified in the National Cultural Heritage - Contributions to the Europeana.eu¹⁴

The Creative Commons

Here is a space of no limits. Everybody could make a contribution to the field and the resources are growing in umbers by the day. Slowly, the memory institutions seem to loose the initiative of being true value mediators in the electronic universe. There are no signs in Romania that this growing output will ever by taken into consideration in the sense of being added and organically integrated in the existing cultural heritage existing context. On some extent this is not to be desired being in the same sense not feasible due to the sheer number of particular aggregated resources.

There are efforts to keep up with the incoming electronic resources. Recently (2011), the Wiki Loves Monuments contest¹⁵ produced a fairly large amount of digital content in an effort to bring to light the best and the most of the visible heritage. After the 2011 edition ended being won by a Romanian, many voices of the civil society addressed the Ministry of Culture asking that all these valuable and freely available resources to be taken into consideration. The outcome was that a part of them were taken by CIMEC as national aggregator for Europeana in order to be integrated in the European collections.

There is also another case of good citizenship initiative that is becoming a beacon for good practices: Encyclopedia of Romania¹⁶ - “enciclopediaromaniei.ro is a non-profit project ran by a team of volunteers who aim to write an electronic, on-line encyclopedia of Romania”.

Arriving to the level of community driven efforts, there are two examples worth mentioning:

- geo-spatial.org: An elegant place for sharing geoKnowledge & geoData, and
- kosson.ro: a community node based on web services addressed to all Romanian specialists in the field of information science (librarians, archivists, museum curators, etc.), eGovernment specialists (information management departments), students willing to learn about the new techniques and technologies for management and data communication, and for all those interested in the field.

All these efforts are leading to the establishment of a new type of Academia, one that is driven by the need to have access to data and information in a timely and easy to apprehend manner building further the existing knowledge forming context.

The European aspirations

The Data

The bases of the knowledge society growth are laid through data accessing. Free access to data is seen as a key enabler for a better coordination of actions in any field that would lead eventually to a more

¹² <http://www.brukenthal-digital.ro/>

¹³ <http://www.cimec.ro/Default.html>

¹⁴ <http://europeana.cimec.ro/>

¹⁵ <http://www.wikilovesmonuments.eu/>

¹⁶ http://enciclopediaromaniei.ro/wiki/Enciclopedia_Rom%C3%A2niei

informed citizen being able to access faster the societal benefits. This data comes from Public Sector as a direct result for getting their activities more transparent and articulated better in a circle or trust established on behalf of, and with the taxpayers.

The next step after changing the habit of "hugging the data"¹⁷ it is to make it link to other bits of data stepping up to the level of Open Linked Data.

In 2011, the W3C Library Linked Data Incubator Group¹⁸ published a Report¹⁹ that explored the state of the art in library practices concerning the Open Linked Data.

Key recommendations of the report are:

- That library leaders identify sets of data as possible candidates for early exposure as Linked Data and foster a discussion about Open Data and rights;
- That library standards bodies increase library participation in Semantic Web standardization, develop library data standards that are compatible with Linked Data, and disseminate best-practice design patterns tailored to library Linked Data;
- That data and systems designers design enhanced user services based on Linked Data capabilities, create URIs for the items in library datasets, develop policies for managing RDF vocabularies and their URIs, and express library data by re-using or mapping to existing Linked Data vocabularies;
- That librarians and archivists preserve Linked Data element sets and value vocabularies and apply library experience in curation and long-term preservation to Linked Data datasets.

Most of the memory institutions are financed through state budget funding and the most valuable of their products are the data they produce, let that be catalogs, statistics or any other forms of raw data aggregations. This particular type of Public Sector Information should find it's way to those who paid for it in forms to their liking as all the memory institutions claim their user-centric approach for their services.

On 17th of November 2003, the European Commission, The European Parliament and The European Council issued Directive 98 "on the re-use of public sector information". In Chapter 1 - General Provisions, Article 1 – Subject matter and Scope it is strictly mentioned that this Directive **shall not apply to**:

- (e) documents held by educational and research establishments, such as schools, universities, archives, libraries and research facilities including, where relevant, organisations established for the transfer of research results;
- (f) documents held by cultural establishments, such as museums, libraries, archives, orchestras, operas, ballets and theatres.

Things evolved and at this moment we speak of the Proposal for a Directive of the European Parliament and of the Council Amending *Directive 2003/98/EC*²⁰ on re-use of public sector

¹⁷ Tim Berners-Lee, The Next Web of Open, linked data, Recorded at TED2009, February 2009 in Long Beach, California. Duration: 16:23, http://blog.ted.com/2009/03/13/tim_berniers_lee_web/

¹⁸ The mission of the group as it was charted in May 2010: "to help increase global interoperability of library data on the Web, by bringing together people involved in Semantic Web activities — focusing on Linked Data — in the library community and beyond, building on existing initiatives, and identifying collaboration tracks for the future."

¹⁹ Library Linked Data Incubator Group Final Report, 2009, W3C, <http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/>

²⁰ http://ec.europa.eu/information_society/policy/psi/docs/pdfs/directive_proposal/2012/proposal_directive.pdf

information that is vividly disputed among the stakeholders. In the Recitals point (7) it is clearly mentioned the area in which this Directive is effective with respect to the works protected by copyright.

The news comes from point (10) and (11) that state that the libraries (including university libraries), museums and archives are covered this time around, and that “to facilitate re-use, public sector bodies should make documents available through machine readable formats and together with their metadata where possible and appropriate, in a format that ensures interoperability [...]”.

This is not the final version on the future Directive, but as it is shown it does offer some indication in which direction the attention should be focused for the memory institution stakeholders.

A surprisingly good practice example comes for Romanian open communities as well. In 2011, “Better Data Award” was given to the *eHarta Historical Maps*²¹, a geo-spatial.ro community effort aimed to digitally preserve valuable Romanian cartographic documents. This award is part of the bigger “Open Data Challenge”²² European competition. This example should be one to follow also for the Romanian LAM community.

The flagship of data sharing initiative in our field offers the best inspiration source: data.europeana.eu. Currently contains open metadata on 2.4 million texts, images, videos and sounds gathered by Europeana.

All these facts lead to further need to realize a national context favorable to data exchange and reuse.

The Information

The important question raised at this point is: what is the starting point for establishing a national data driven context?

First we go back to the milestone set in 2008: the Order of the Ministry of Culture²³ to build a national portal for cultural heritage digital resources – Culturalia. This initiative in this point in time is backed strongly by the electronic resource aggregation “Mobile Cultural Objects Classified in The National Cultural Heritage” – a database feeding the Europeana.eu with almost 35000 items.

Having missed the objective of building a national common infrastructure up to this point, other alternatives are to be considered given the fact that some of the big players in terms of the size of their digital collections are concerned.

These resources providers should come to a common level of sharing their metadata in a concentrated single point of access for the aggregation. The easiest way to come to this much-desired milestone is harvesting into a single point via OAI-PMH²⁴ the repositories. Further down there is a figure presenting the biggest aggregations of cultural heritage digital objects existing right now. The stakeholders were picked from many other smaller providers for the out-of-the-box capacity to share information via a common standard provided by their individual software solutions.

It is easy to go further and envision a single point of access where one interested user could make use of the resources benefiting from a federated repository infrastructure.

²¹ <http://earth.unibuc.ro/articole/eHarta?lang=en>

²² <http://opendatachallenge.org/>

²³ Order of the Minister of Culture No. 2467/2008 - <http://poliptic.wordpress.com/2012/06/03/culturalia-un-portal-romanesc/>

²⁴ Open Archives Initiative – Protocol for Metadata Harvesting <http://www.openarchives.org/pmh/>

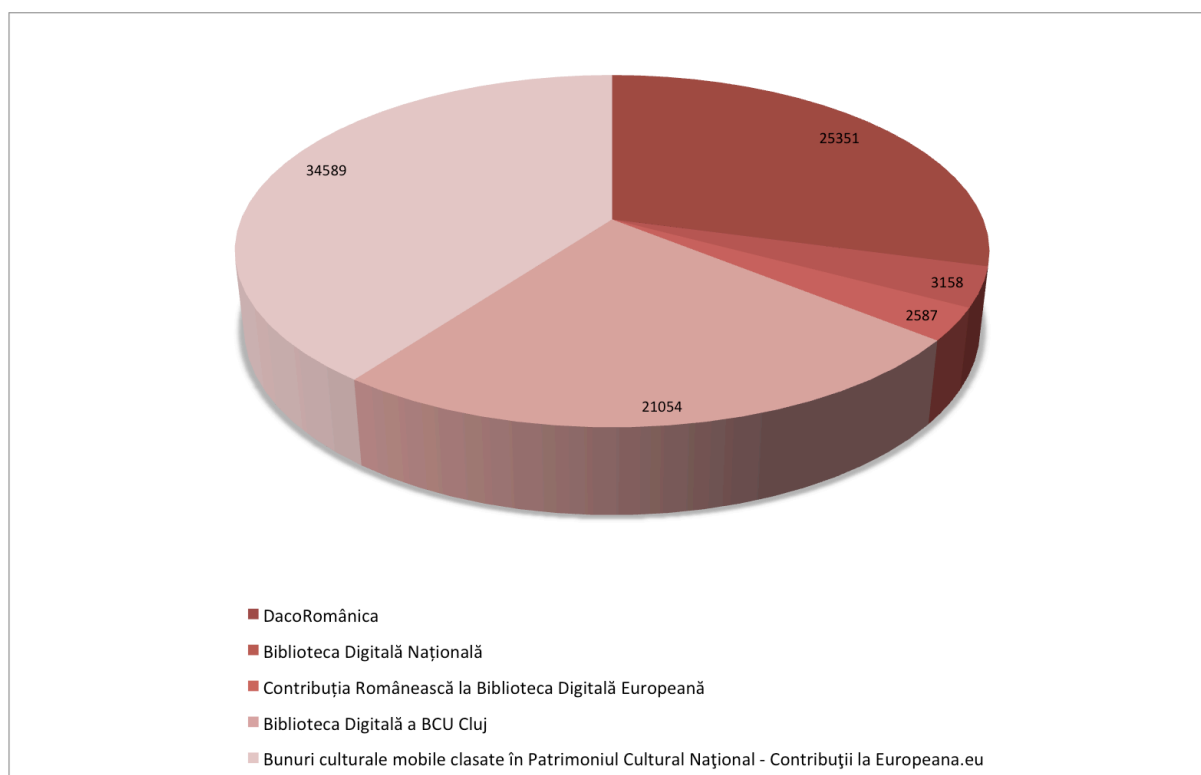


Figure 3: Highly visible stakeholders and the collections: Romanian Digital Core

This move would set an example for the next level of aggregation for all existing resources, many who are not benefiting from systems capable of making them talk one to each other – specialized repository solutions. This type of easy traveling information across collections would ensure firstly an easy level for further expansion and a better federation and with other similar local, European or international projects like Culturalia.ro, Europeana.eu, Digital Library of America or World Digital Library.

A single point for access to this, let's call it Romanian Digital Core (RDC), would suffice the need for raw data as well. RDC should publish metadata of all the aggregated collections as Open Data.

Undoubtedly, this should be the very first step because the rest of the key players on the scene of Romanian information brokerage are needed in order to complete the *Virtuous Circle of the digital economy*²⁵.

The Context

The starting point for realizing a rich knowledge context could be possible once the Romanian Digital Core will be set in place. This move should be accompanied by the integration of the library catalogs provided by the Romanian library system.

One first milestone would be realized through the integration of [ROLiNeST](#) federated search service in tandem with [TinREAD](#) county libraries federated search service. If such an approach would prove to be cumbersome in terms of technical barriers, a campaign for linking each and every catalogue should take its place.

²⁵ A Digital Agenda for Europe - http://ec.europa.eu/information_society/digital-agenda/documents/digital-agenda-communication-en.pdf

This milestone would prove highly beneficial for all involved stakeholders, as it will offer a second advantage: data normalization. In truth it will achieve another secondary goal: the national shared catalogue in one sweep.

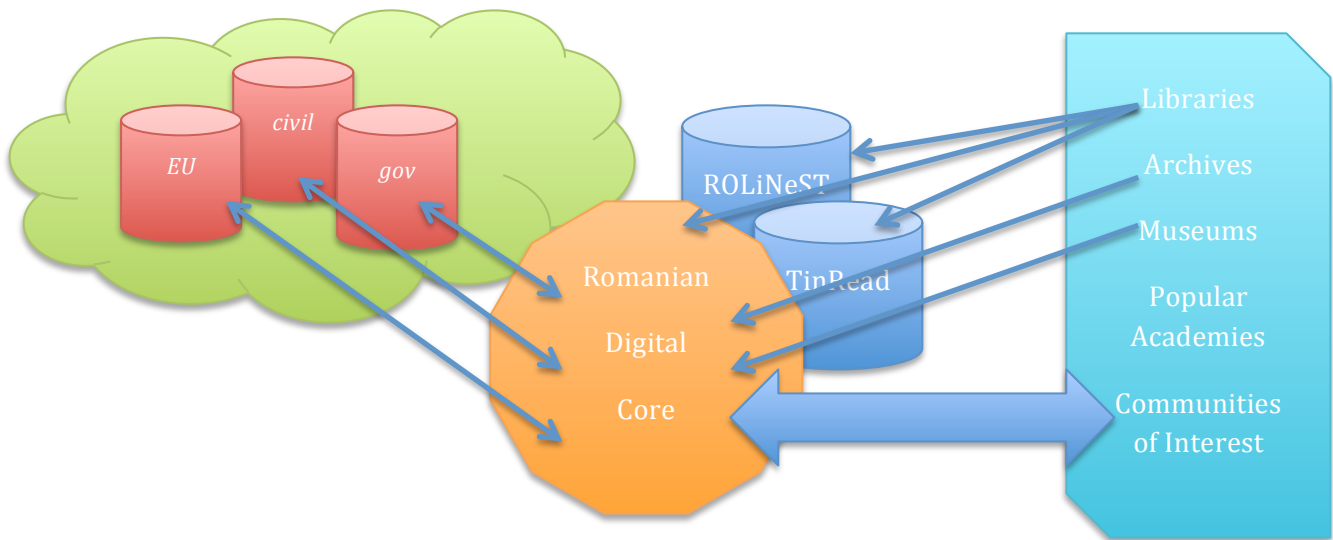


Figura 4: A possible scenario for the future context

Achieving such a knowledge-enabling context is not an easy task. The good news is that as long as the core is functioning proving incentives for further developments, the rest of the scenario will unfold in concentric waves driven by easy to use technical means for communication and exchange of data. The data, this is the key enabler for achieving such a type of integration. As long as a single point for aggregation will exist, the data added continuously will deliver many and unexpected connections for the users. The users should be allowed to re-use the data and to further enrich it in a true semantic driven environment.